MIPT Data Visualization Course

Data Visualization in Modern Machine Learning

Ashuha Arseniy[1,2]

Bayesian Research Group[1], MIPT[2]

ars.ashuha@gmail.com

November 9, 2016

Motivation

Motivation

- It's a little bit sad, but we can plot only 2D data, isn't it?
- A goal of data visualization is to understand of inner data structure.
- Or represent data in much more interpretable form.

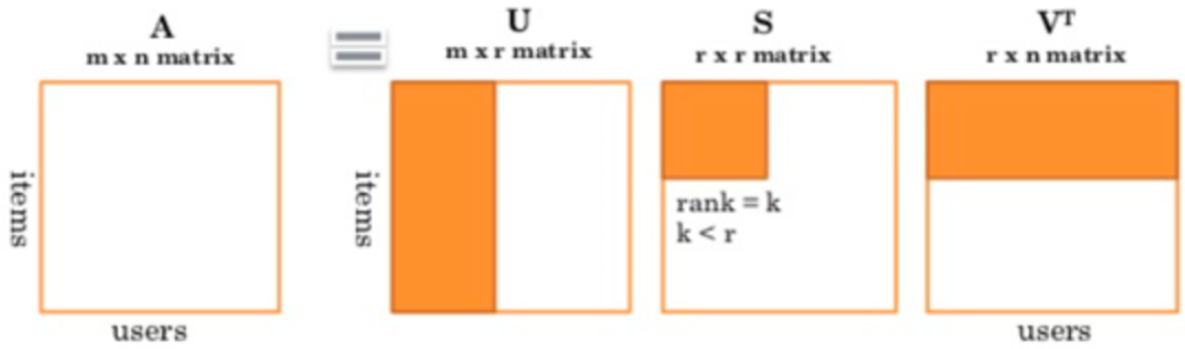
Two way to get this goal:

- Low Rank Way (SVD, Auto-encoders, LDA, etc.)
- Generative Models Way (GAN, Image Capturing, etc.)

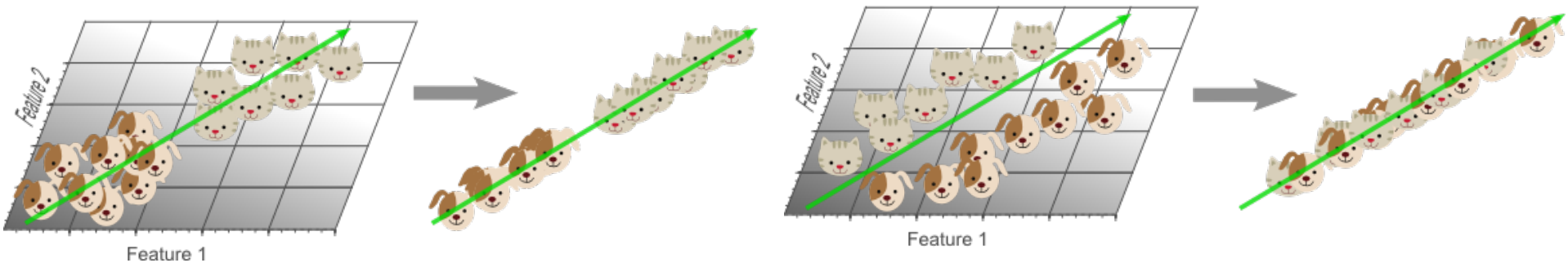
Low Rank idea

- We have matrix $X_{items \times features}$
- Let's try to represent each item's vector as a smaller dimension ones
- What should we do?

Principal component analysis is a matrix decomposition, minimize L_2 norm

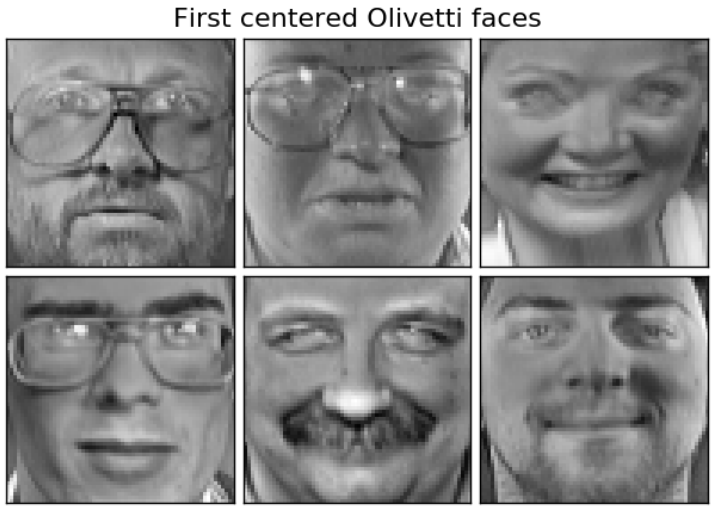


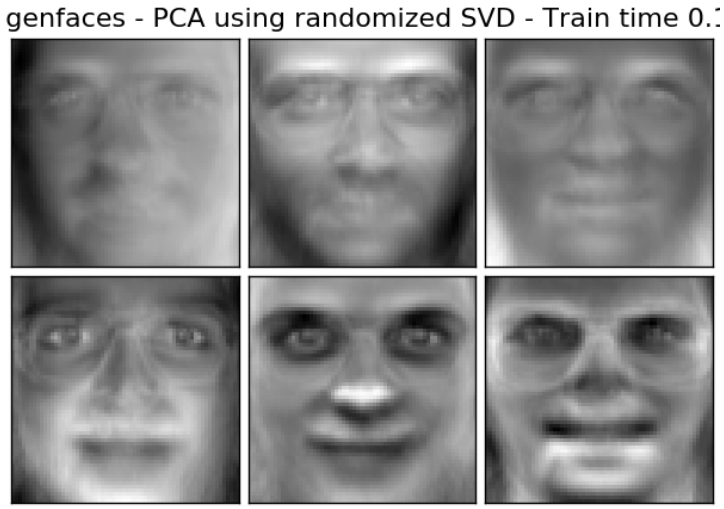
Intuition save maximum data variance decomposition, minimize L_2 norm



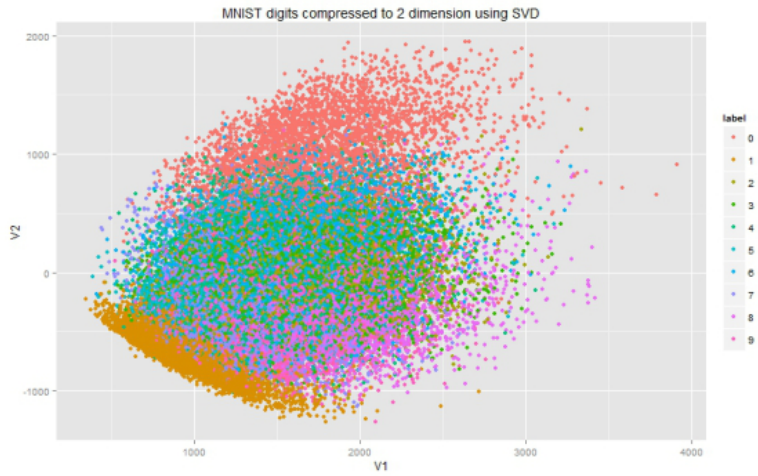
SVD: Faces dataset

Main components:



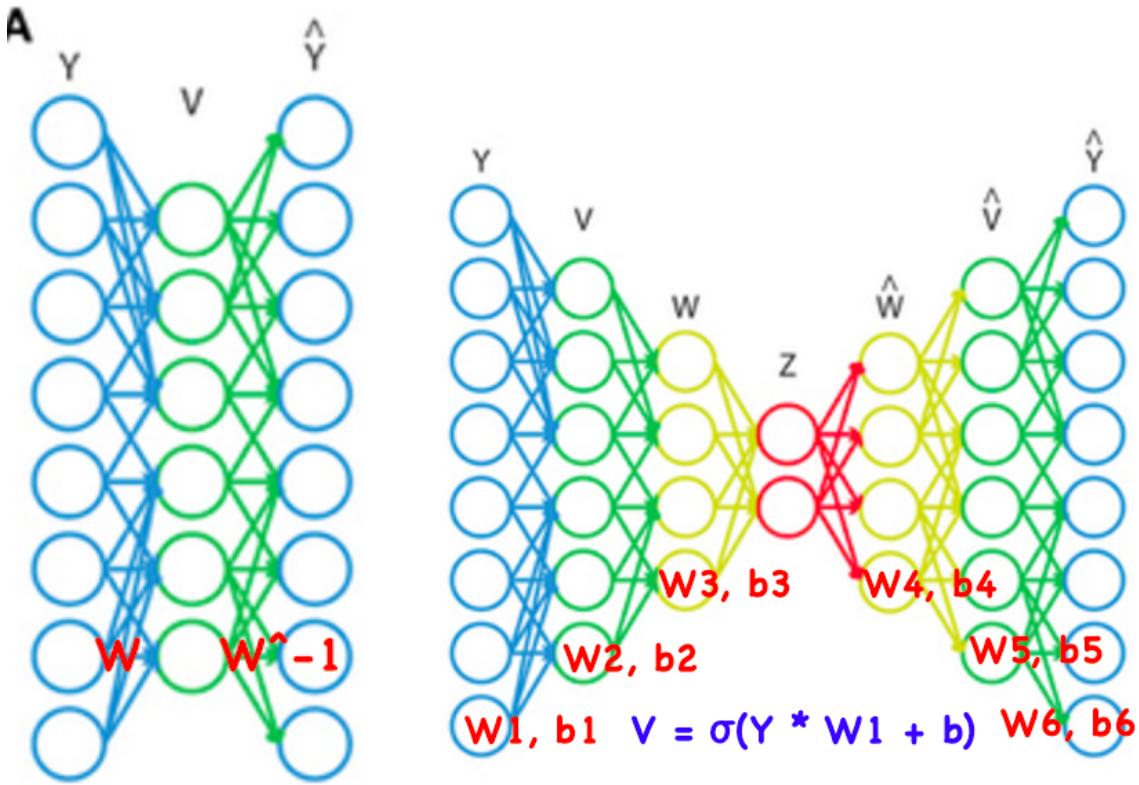


Plot in 2d:



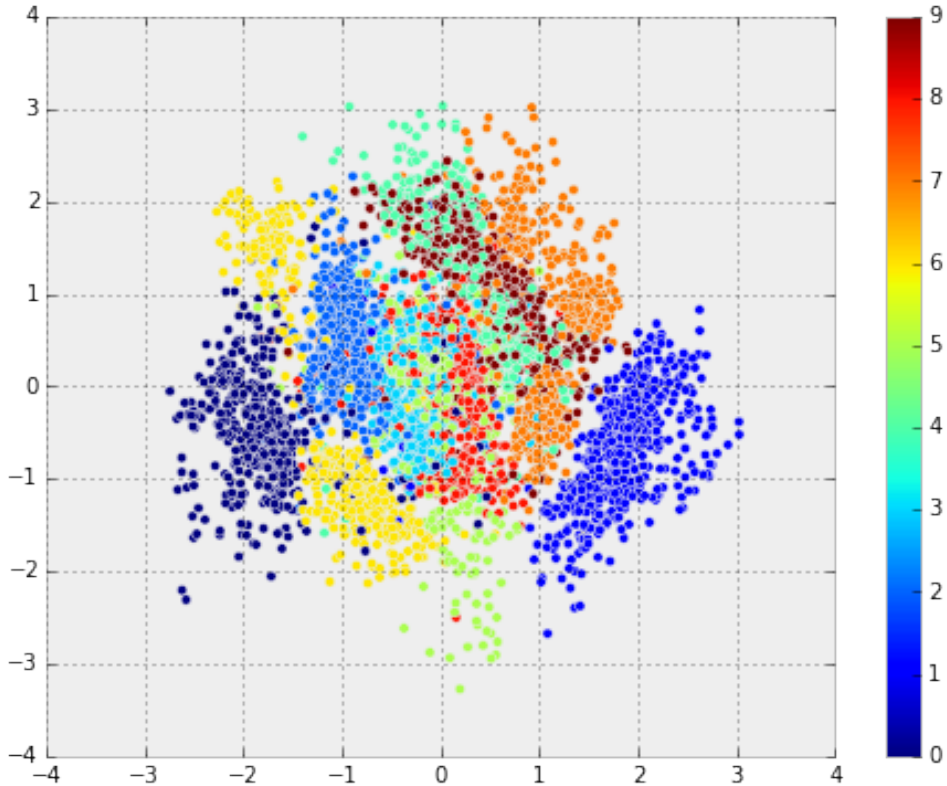
Non-linear Expansion

- What did we do wrong? Our picture mix different classes and so on.
- Let's try non-linear generalization.



- How to find W_n?
- Define loss function $L(Y, \hat{Y})$ and use your favourite opt method.

Auto encoders example



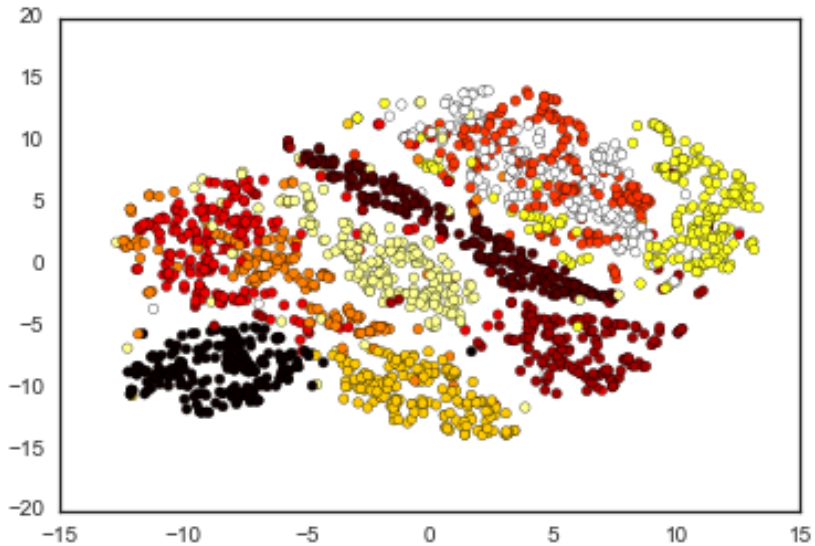
http://dpkingma.com/sgvb_mnist_demo/demo.html

Stochastic Neighbor Embedding

X – high dimensional obj and Y – low dimensional ones, σ – width params

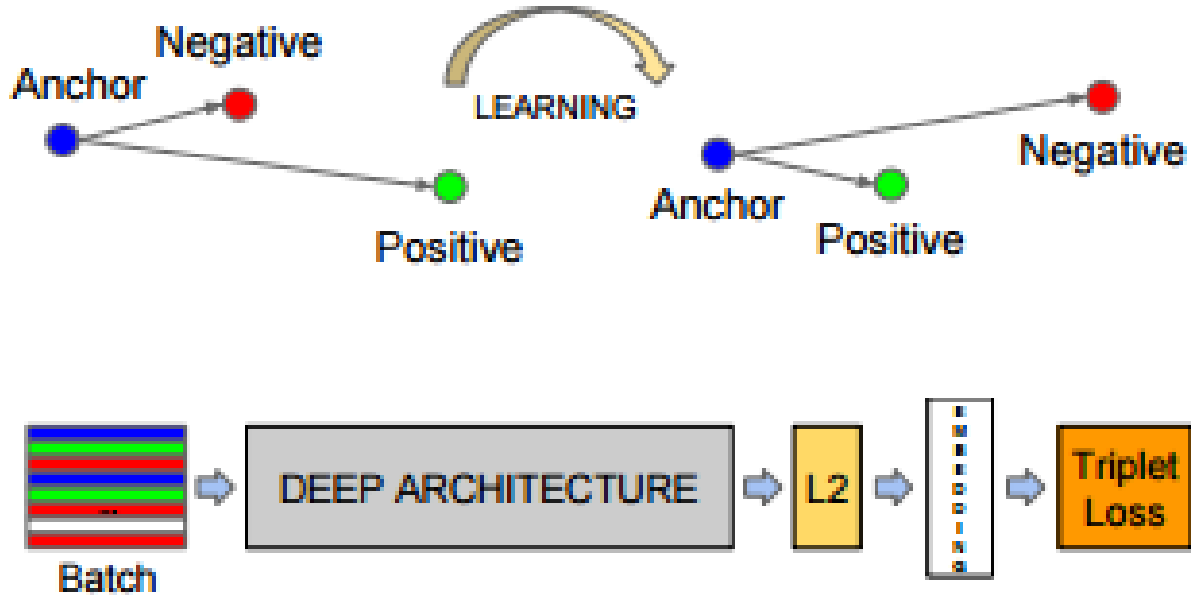
$$p_{j|i} = \frac{\exp(-\|\mathbf{x}_i - \mathbf{x}_j\|^2/2\sigma_i^2)}{\sum_{k\neq i}\exp(-\|\mathbf{x}_i - \mathbf{x}_k\|^2/2\sigma_i^2)} \qquad q_{j|i} = \frac{(-\|\mathbf{y}_i - \mathbf{y}_j\|^2)}{\sum_{k\neq i}(-\|\mathbf{y}_i - \mathbf{y}_k\|^2)}$$

$$KL(P||Q) = \sum_j \sum_i p_{i|j} \log \frac{p_{i|j}}{q_{i|j}} \to \min_q$$



Deep Neural Nets + t-SNE (modification of SNE with Student test):
http://cs.stanford.edu/people/karpathy/cnnembed/

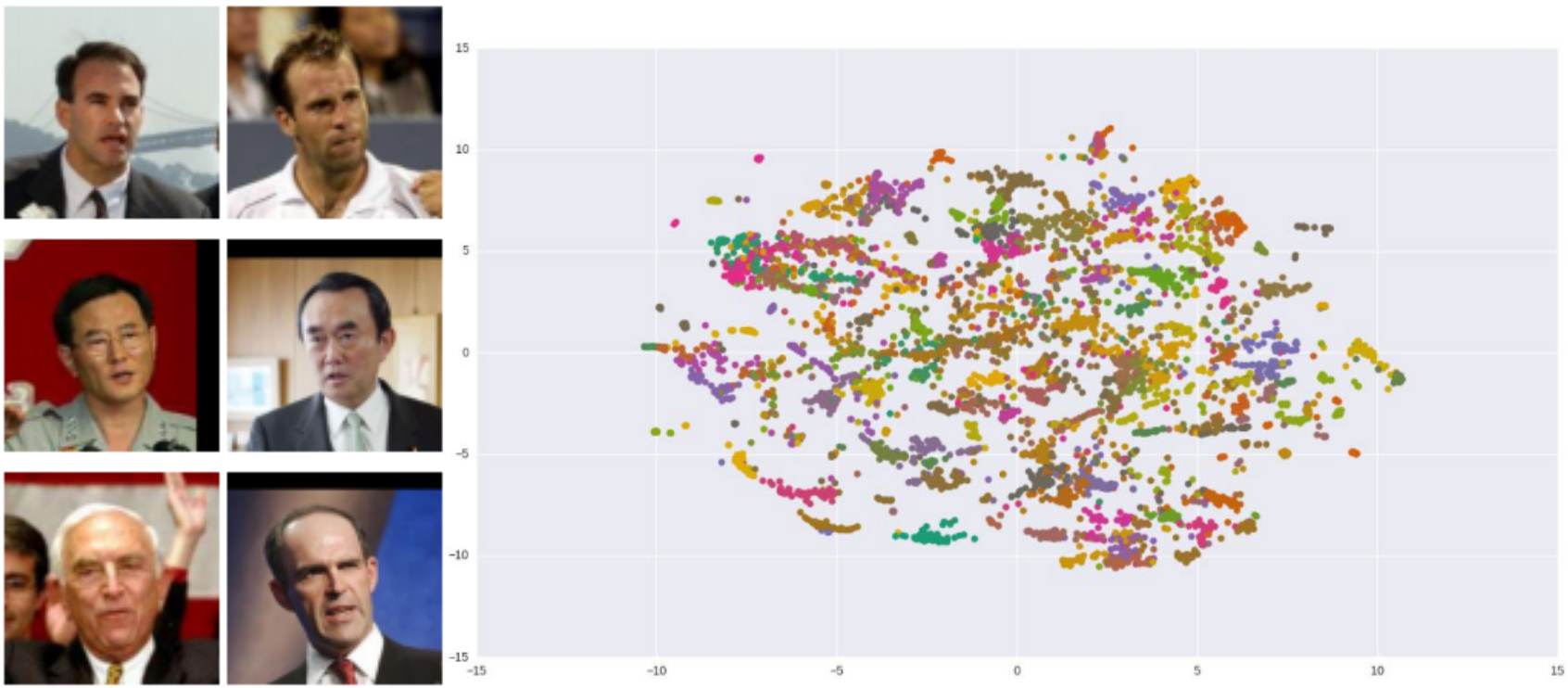
DNN Metric Learning Triplet Loss



The loss that is being minimized is then

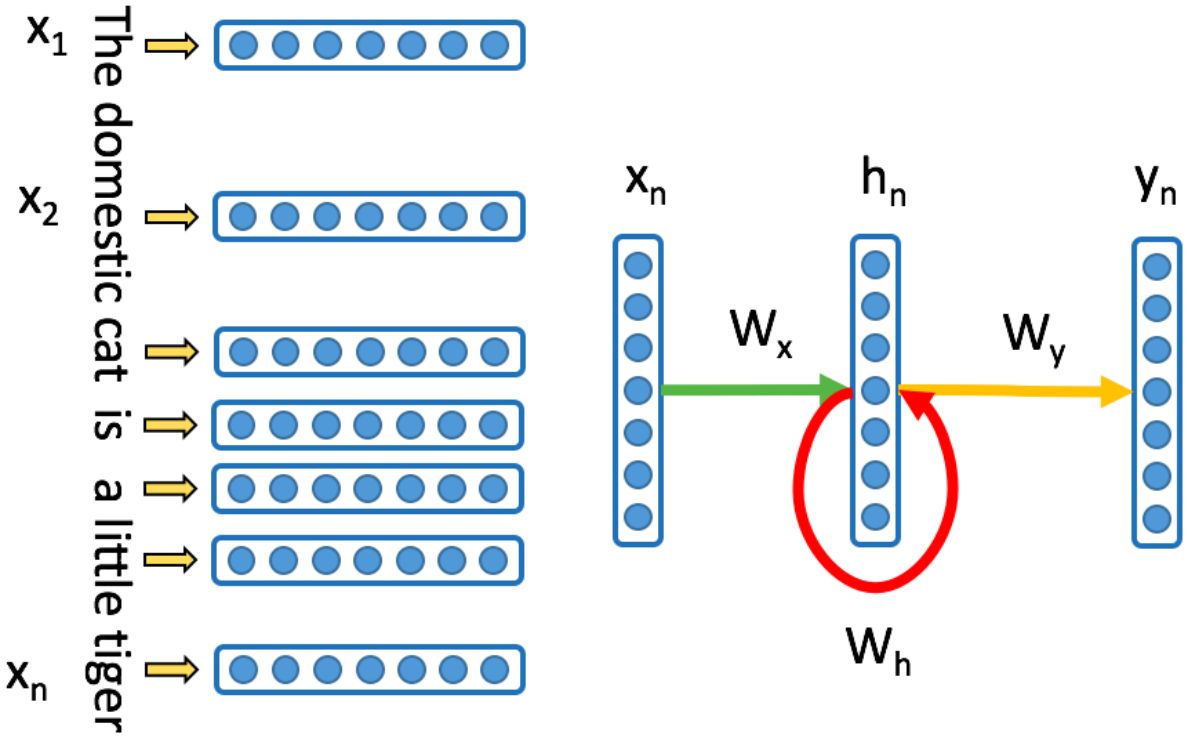
$$\sum_{i}^{N} \left[\|f(x_i^a) - f(x_i^p)\|_2^2 - \|f(x_i^a) - f(x_i^n)\|_2^2 + \alpha \right]_+$$

DNN Metric Learning Triplet Face and Music



High level

- We have mapped each object into vector
- Let's train this vector for match complex object like words

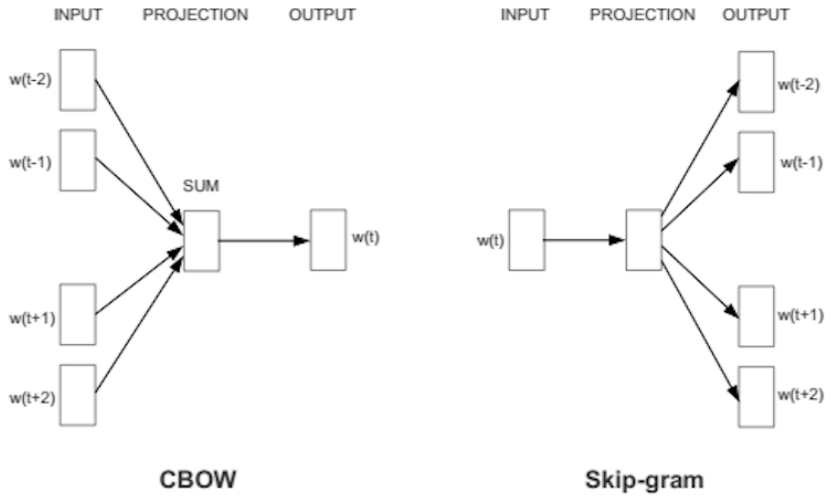


$$h_n = W_x x_n + W_h \sigma(h_{n-1})$$

$$y_n = \sigma(W_y h_n)$$

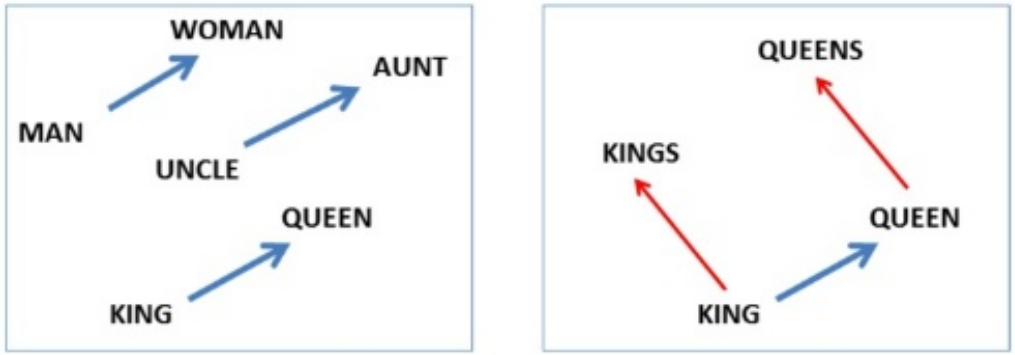
Word2Vec

- Shallow Neural Net



- Operations on embeddings are great

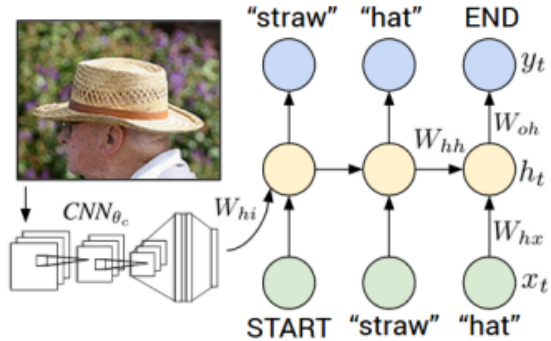
vec("man") – vec("king") + vec("woman") = vec("queen")



http://turbomaze.github.io/word2vecjson/

Image2Text

- Ok, we have a picture and want to represent in lower dim space
- lets try map picture om a word sentence space



man in black shirt is playing guitar.

construction worker in orange safety vest is working on road.

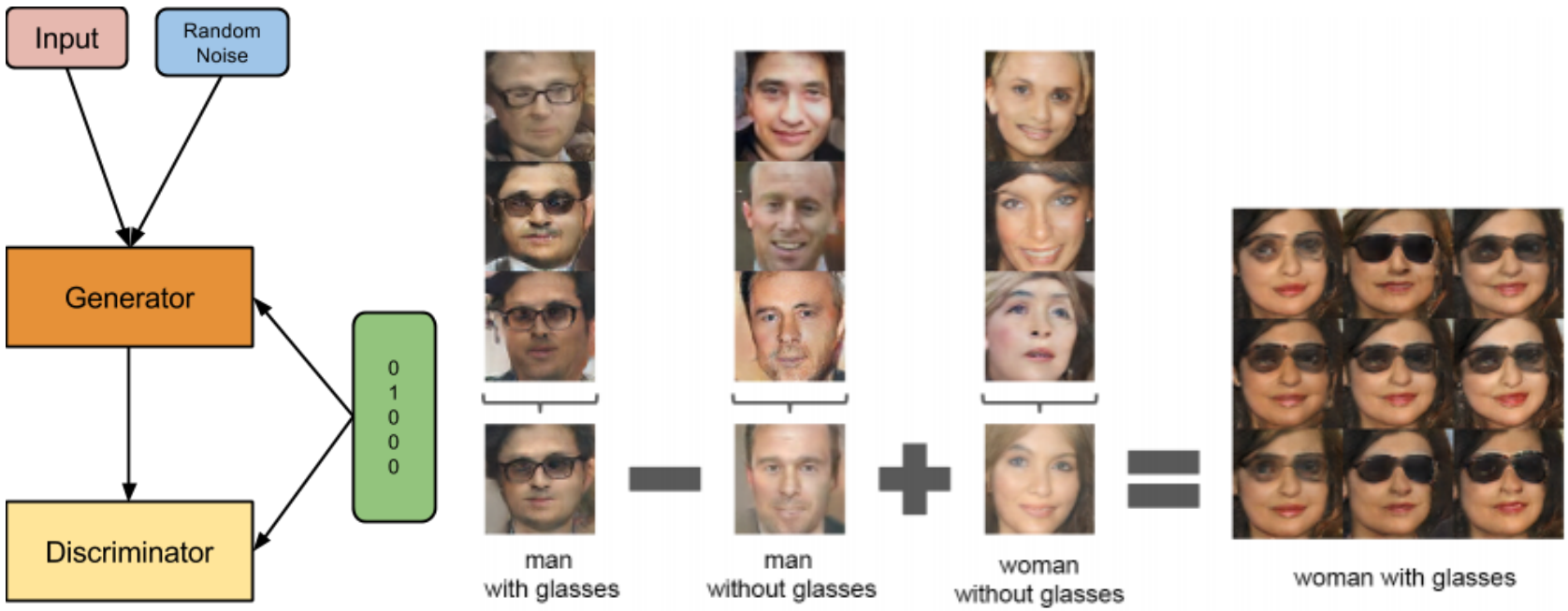
two young girls are playing with lego toy.

boy is doing backflip on wakeboard.

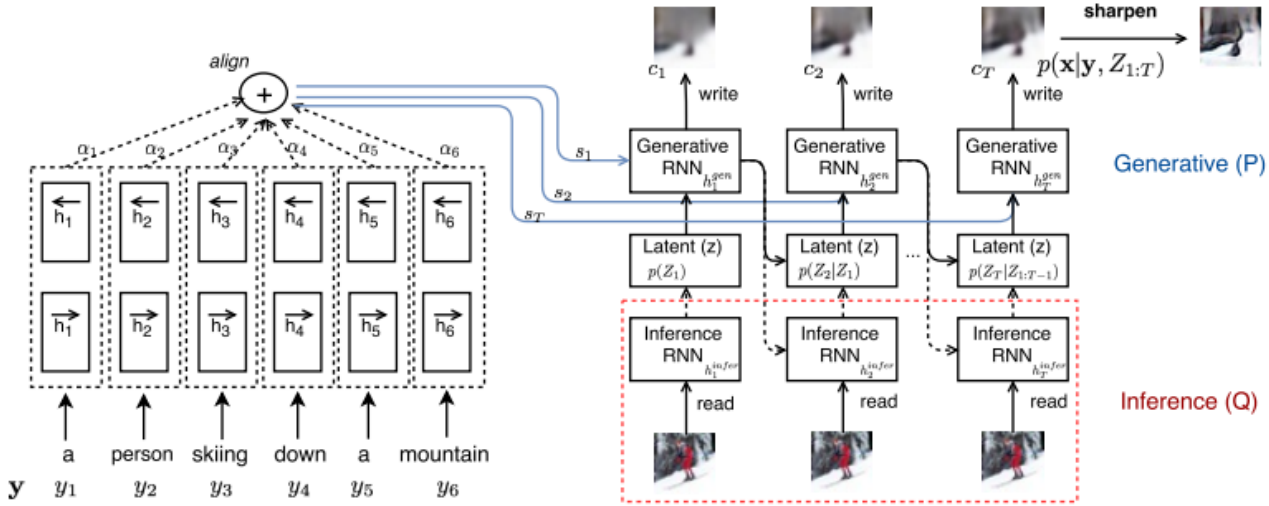
http://cs.stanford.edu/people/karpathy/deepimagesent/
http://cs.stanford.edu/people/karpathy/deepimagesent/rankingdemo/

Generative Adversarial Networks

- Image Generation is a lintel bit hardcore
- Most modern idea is like this



Text2Image



A stop sign is flying in blue skies.

A herd of elephants flying in the blue skies.

A toilet seat sits open in the grass field.

A person skiing on sand clad vast desert.

https://arxiv.org/pdf/1511.02793v2.pdf

References